

---

# Top Mass Measurements at CDF

---

BEACH 2004, Chicago, July 2



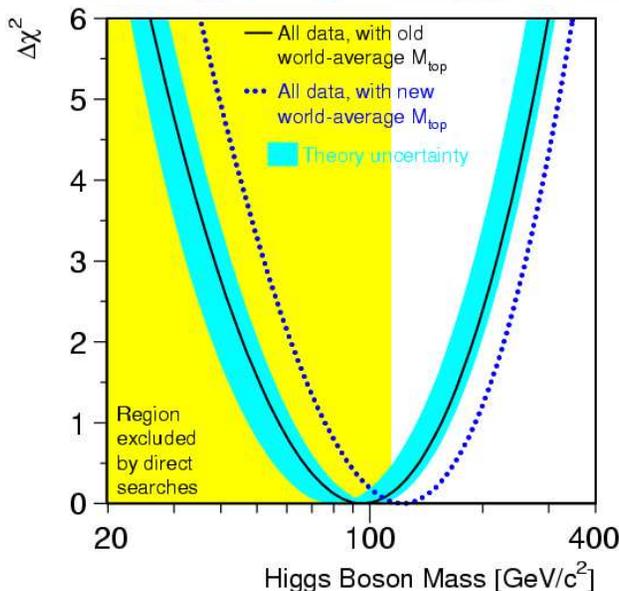
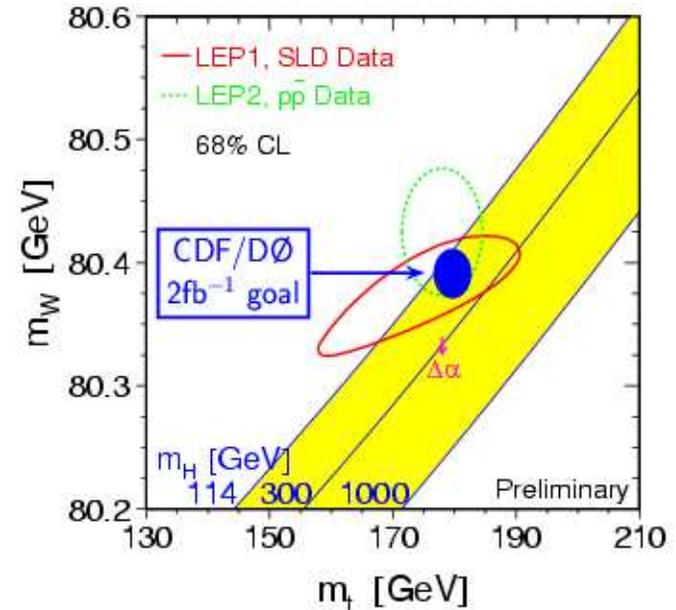
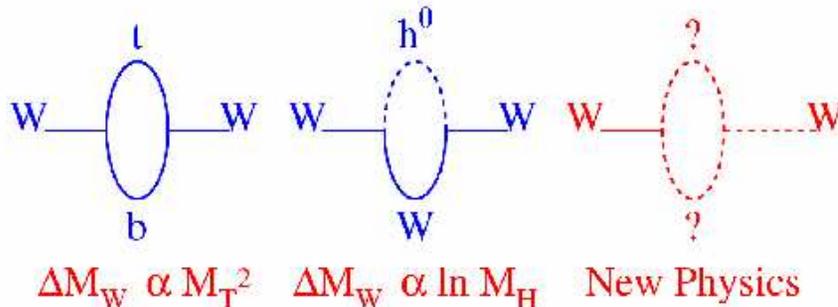
Pedro Movilla Fernández  
(LBNL)  
for the CDF Collaboration



# Why Measure the Top Mass?



- Fundamental parameter
- Correlated to other SM parameters via electroweak corrections



Updated  $m_t$  world average:  
 [Nature 429 (2004) 638]  
 $m_t = 178.0 \pm 4.3 \text{ GeV}/c^2$  (old:  $174.3 \pm 5.1 \text{ GeV}/c^2$ )

Updated constraint on Higgs mass:  
 $m_H = 113^{+62}_{-42} \text{ GeV}/c^2$  (old:  $96^{+60}_{-38} \text{ GeV}/c^2$ )  
 $m_H < 237 \text{ GeV}/c^2$  (old:  $m_H < 219 \text{ GeV}/c^2$ )

**Precise measurement  
 provides stringent SM tests**



# SM Top Quark Signatures



## Top Production (Tevatron):

mainly in pairs via

**85%  $q\bar{q}$  annihilation**

15% gg fusion

$\sigma_{t\bar{t}}(1.96\text{TeV}) \sim 6.7\text{pb}$  (theory)

## Top Decay:

$\tau \sim 10^{-24}\text{sec}$

$t \rightarrow Wb$

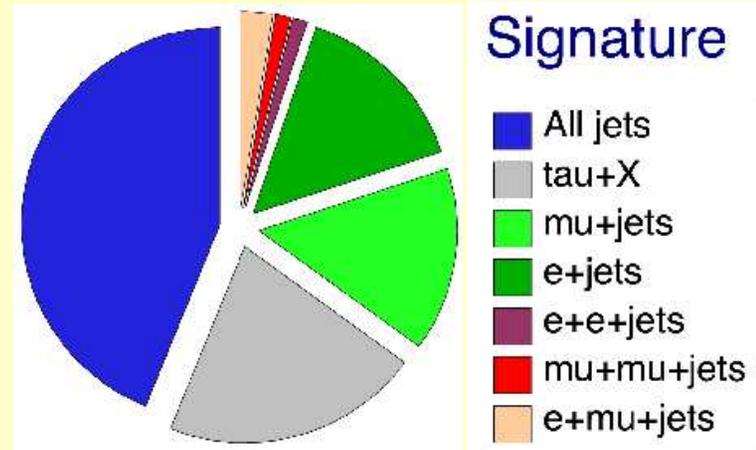
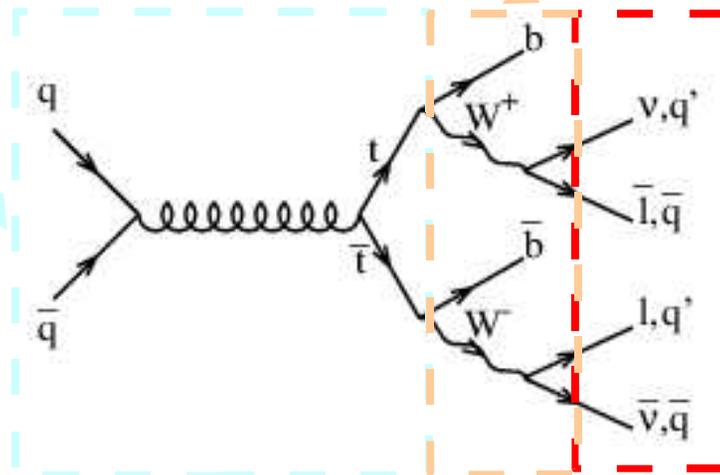
BR  $\sim 99.9\%$

## W Decay:

defines top event signature

$W \rightarrow q\bar{q}'$  hadronic (BR=6/9)

$W \rightarrow l\nu$  leptonic (BR=3/9)



## General Event Topology:

- spherical events ( $t\bar{t}$  production near threshold)
- 2 b jets from top quarks (crucial for ident.)**
- jets/leptons with high transverse energy  $E_T$
- large missing  $E_T$  in leptonic decay modes

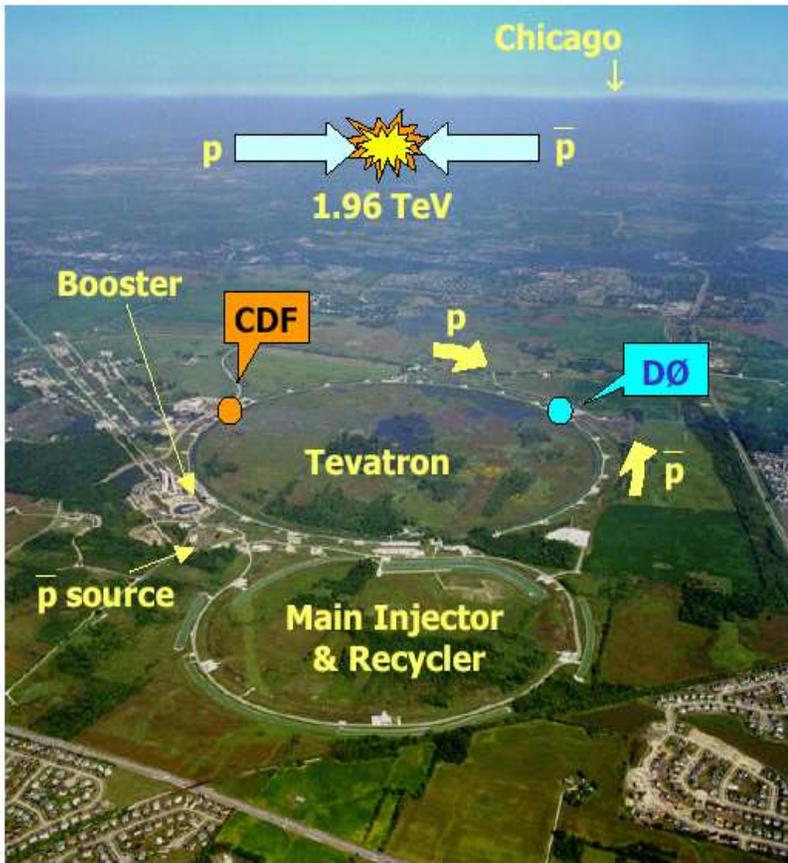
## CDF II Top Mass Measurements:

Dilepton: 2  $e/\mu$ , 2 jets, large miss.  $E_T$   
(BR=5%, S/B $\sim$ 10)

Lepton+Jets: 1  $e/\mu$ , 4 jets, large miss.  $E_T$   
(BR=30%, S/B $\sim$ 1)



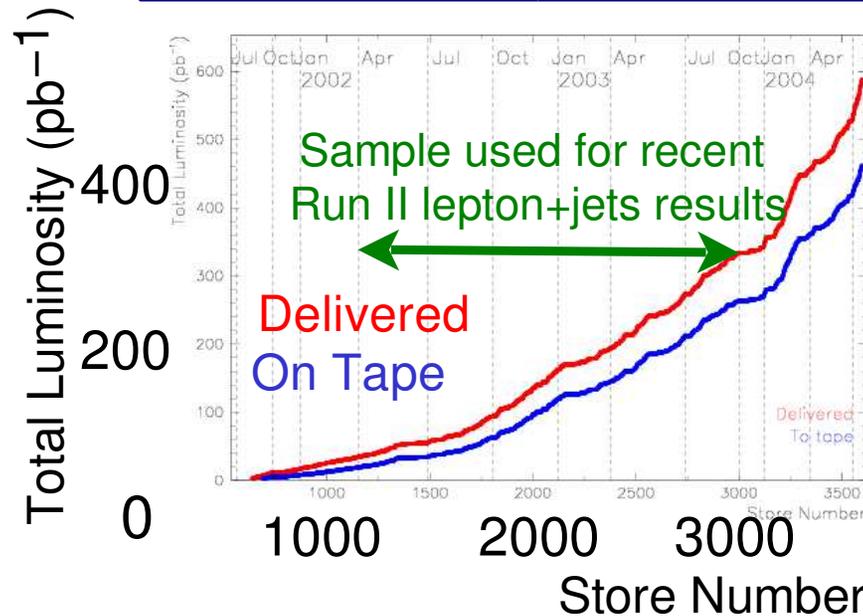
# Tevatron Run II



CM energy 1.96 TeV (Run I: 1.8 TeV)  
Increase of  $t\bar{t}$  cross section by ~30%

Substantial increase of luminosity:  
Record:  $\sim 8 \times 10^{31} \text{ cm}^{-2} \text{ s}^{-1}$

## CDF II integrated luminosity



Total on tape (now):	460 pb <sup>-1</sup>
Analysis samples:	
Dileptons:	126 pb <sup>-1</sup>
Lepton+Jets:	162 pb <sup>-1</sup>
(with at least 1 SVX tag)	



# Top Event Selection (Lepton+Jets)



## CDF II detector:

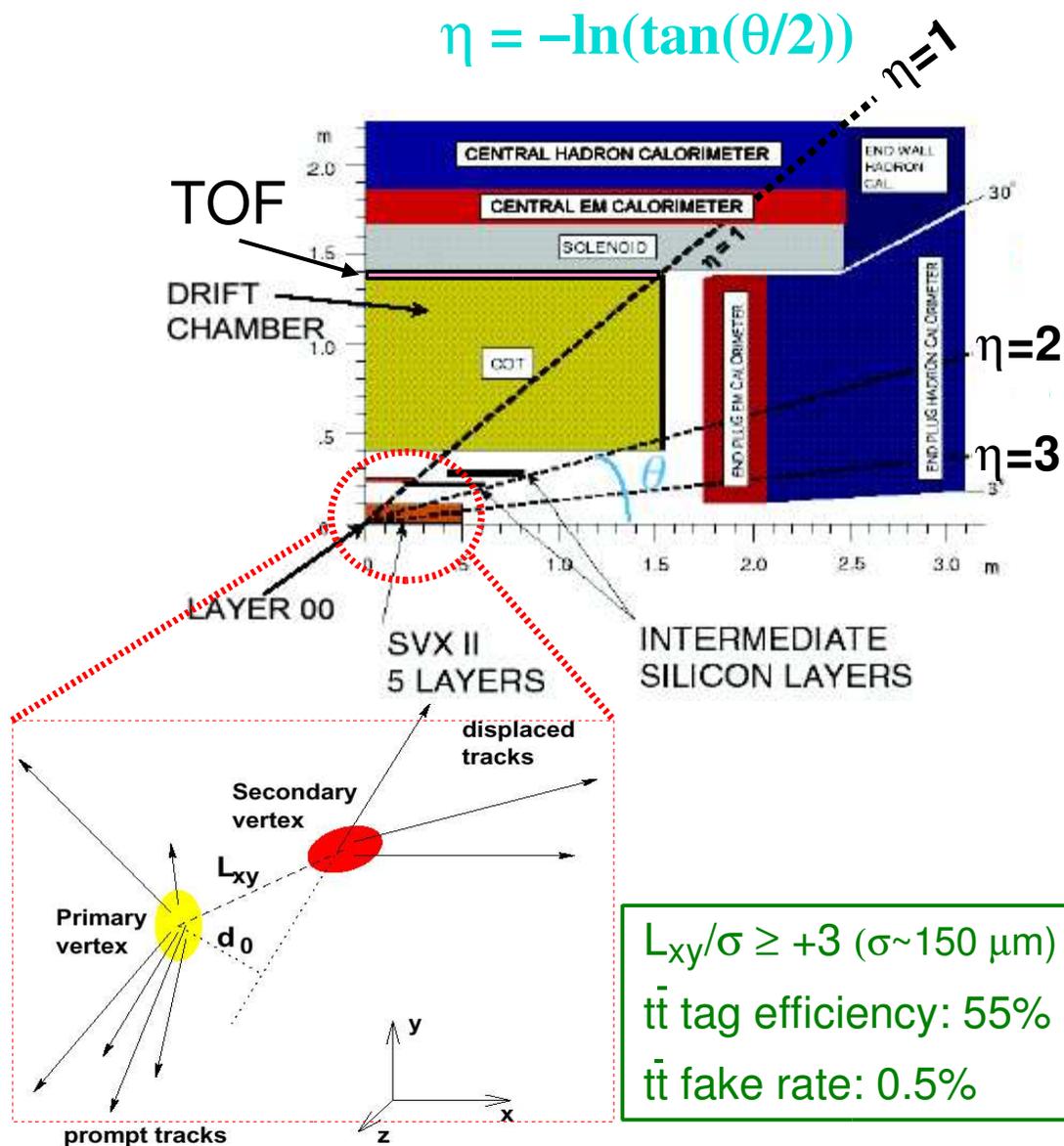
- Calorimeter (EM, HAD)
- Tracking system
- Vertex detector

### Basic kinematical cuts:

1. One Lepton  $E_T > 20$  GeV
2. Missing  $E_T > 20$  GeV
3. Four jets  $E_T > 15$  GeV,  $|\eta| < 2$   
(depend on analysis)
4. At least one SVX tag

### SVX b jet tagging:

Identification of **displaced decay vertices** from long-lived B-hadrons





# Methods



## I. Template Method (**TM**) (= Run I Method):

- Kinematic fitter to reconstruct top mass
- Kinematic constraints
- Use "best" of 12 combinations (4 if double b tag)  
(6 jet-quark combinatorics, 2 neutrino solutions)
- 1-dimensional templates parametrized as function of top mass

## II. Multivariate Template Method (**MTM**):

- Refined kinematic fitter with jet energy scale optimization
- Kinematic constraints
- Use "best" of 12 combinations (4 if double b tag),  
Weight according to correct permutation probability
- Multidimensional non-parametric templates

## III. Dynamical Likelihood Method (**DLM**):

- Matrix element likelihood
- Use all 12 combinations (4 if double tag)
- Use calorimeter transfer functions



# Template Method (TM)



Reconstruct invariant top mass for each event

- Minimize  $\chi^2$  expression, kinematic constraints  
 $m_t = m_{\bar{t}}$ ,  $M_{W^+} = M_{W^-}$ ,  $p_{t\bar{t}}$  balance;

$$\chi^2 = \sum_{\ell, jets} \frac{(\hat{P}_T - P_T)^2}{\sigma_{P_T}^2} + \sum_{i=x,y} \frac{(\hat{U}'_i - U'_i)^2}{\sigma_{U'_i}^2} + \frac{(M_{\ell\nu} - M_W)^2}{\sigma_{M_W}^2} \\ + \frac{(M_{jj} - M_W)^2}{\sigma_{M_W}^2} + \frac{(M_{\ell\nu j} - M_t)^2}{\sigma_{M_t}^2} + \frac{(M_{jjj} - M_t)^2}{\sigma_{M_t}^2}.$$

- Combinatorial problem: 12(4) solutions for 1(2) b tags  
2  $p_z$  neutrino solutions, 6(2) jet-parton combinations;  
(we don't distinguish between q and  $\bar{q}'$  from W decay); **Use smallest  $\chi^2$  solution;**

Build **top mass templates** from MC samples

for signal process with different  $m_t$  and for the background processes;

Calculate top mass likelihood

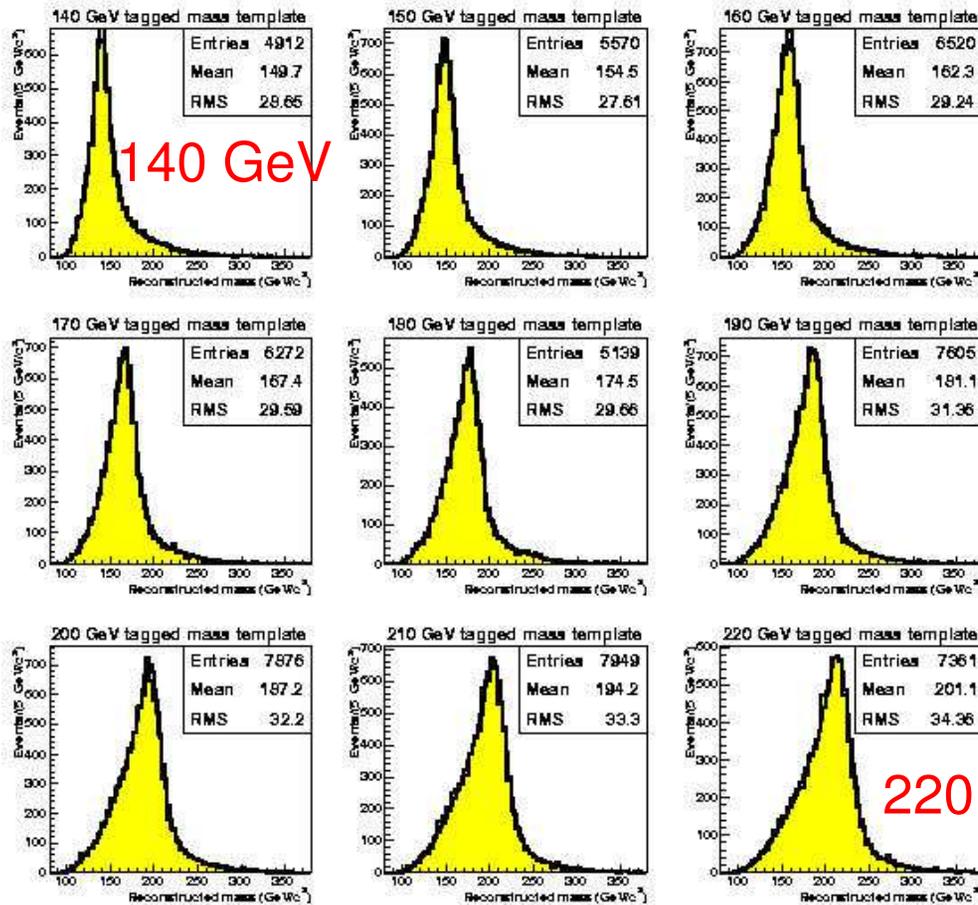
- Use templates as probability densities to be compared with data in order to derive a **top mass probability for each event**
- Top mass is value of  $m_t$  which **maximizes likelihood for the whole data sample**  
= product of event-by-event top mass probabilities (**unbinned likelihood fit**)



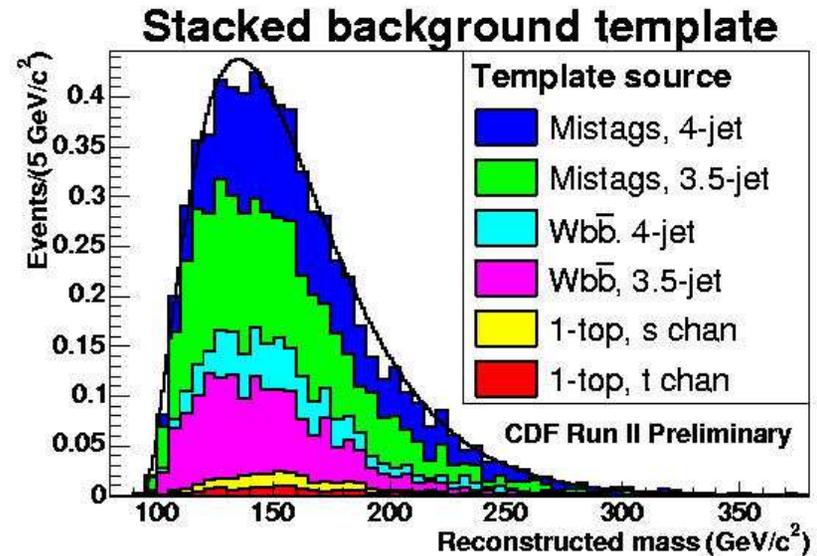
# TM Templates



## CDF Run II Preliminary



Reconstructed invariant top mass



Signal:

Herwig

Background:

Wb $\bar{b}$  (Alpgen+Herwig)

Single Top (Pythia)

Mistags (W4p Alpgen+Herwig)

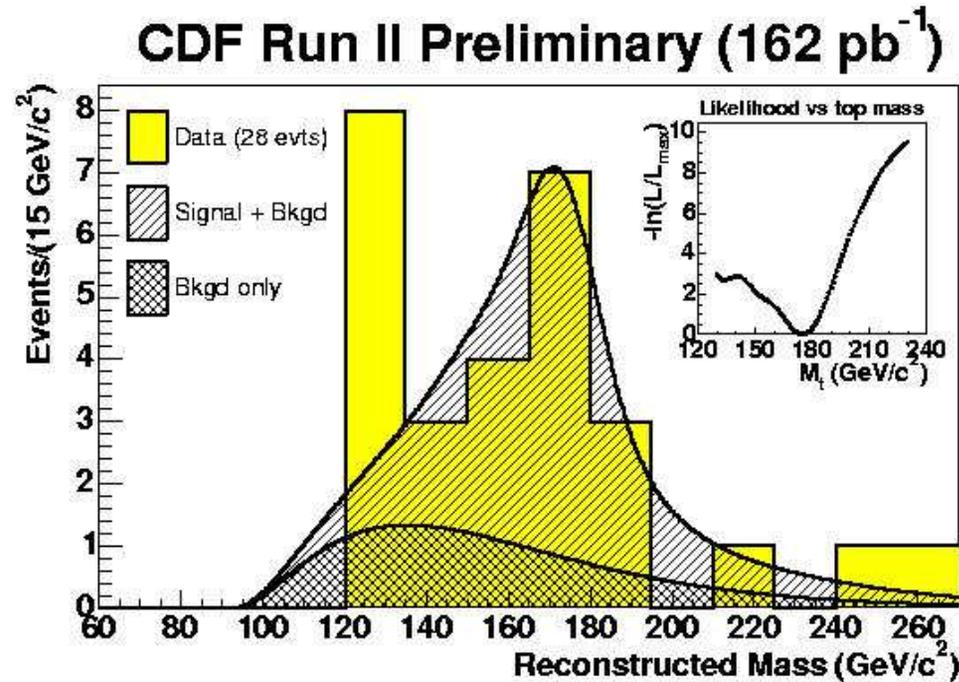
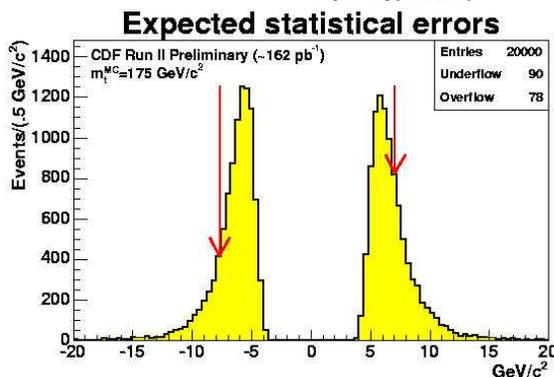
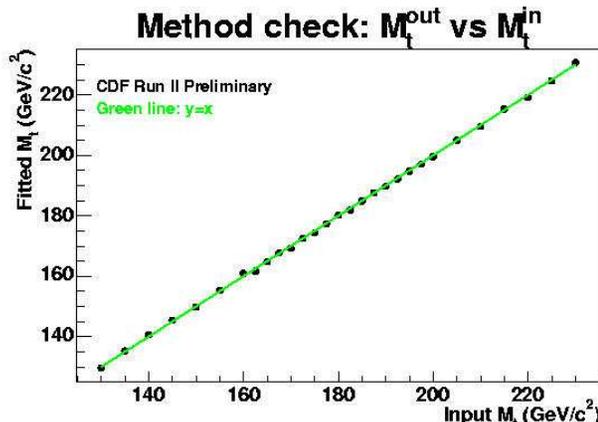
Signal templates are parametrized by continuous functions of  $m_t$



# TM Results



28  $t\bar{t}$  candidate events  
 $6.8 \pm 1.2$  background events  
 (expected value from cross section measurement)  
 ... fixed in the fit



$$M_{\text{top}} = 174.9^{+7.1}_{-7.7} \text{ (stat.)} \pm 6.5 \text{ (syst.) GeV}/c^2$$

Jet energy systematic:  $\pm 6.3 \text{ GeV}/c^2$



# Systematic Uncertainties



**Jet energy systematics are by far dominant**

Jet Energy Scale	6.3
Initial State Radiation	0.4
Final State Radiation	0.9
Parton Distribution Functions	0.2
Generators	0.4
Other MC Modeling	0.7
Background Shape	0.8
B-tagging	0.1
Total	6.5

**Preliminary, will be reduced soon!**

Relative to Central	3.0
Central Calorimeter Response	4.6
Corrections to Hadrons (Absolute Scale)	2.2
Corrections to Partons (Out-of-Cone)	2.3
Total	6.3

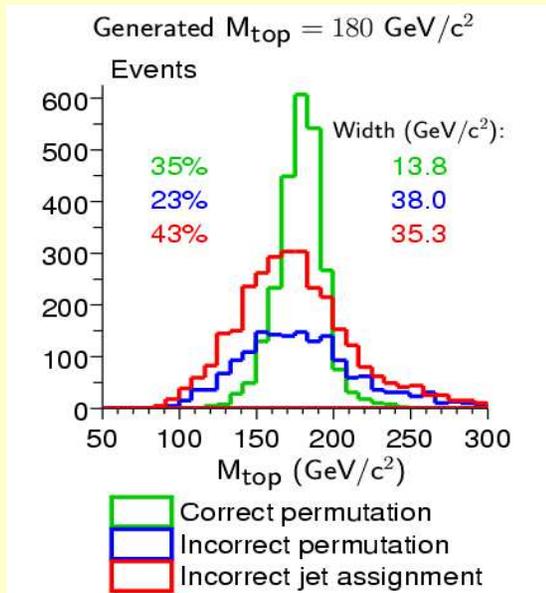
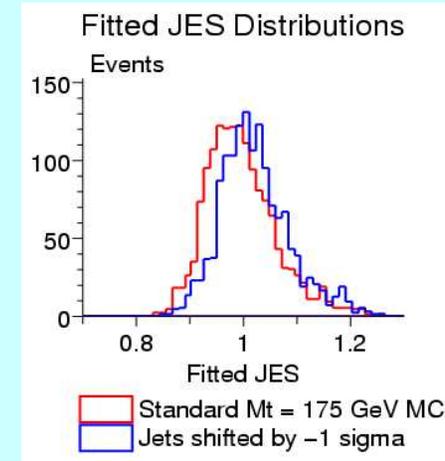


# Multivariate Template Method (MTM)



1) Kinematic fitter for event reconstruction:  
 Includes adjustable jet energy scale factor (JES)  
 to be calibrated in the reconstruction of the  
 $W \rightarrow q\bar{q}'$  decay by using  $W$  mass constraint;  
 Multiply all jet energies with JES scale factor

- ➔ decrease of JES systematics
- ➔ increase of statistical error



2) Three types of signal templates:

- Correct permutation samples (**CP**)
- Incorrect permutation samples (**IP**)
- Incorrect jet assignment samples (**IJ**)

Uses information from fit and from  $t\bar{t}$  production/decay dynamics to predict CP probability and to weight signal templates accordingly. Useful quantities are

- ➔ fit  $\chi^2(\text{permutation } i) - \chi^2(\text{best permutation})$
- ➔  $\cos\angle(\text{lepton, leptonic } b) \text{ in } W \text{ rest frame}$
- ➔  $t\bar{t}$  spin correlation term

”Knowledge” of template type improves mass resolution

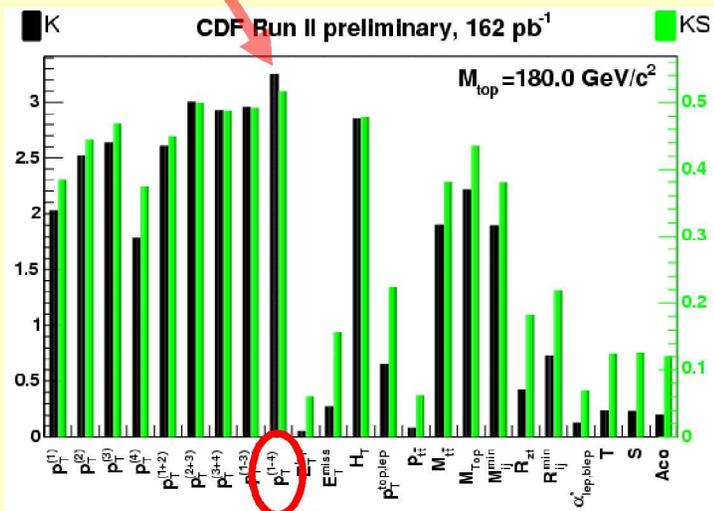


# MTM Templates

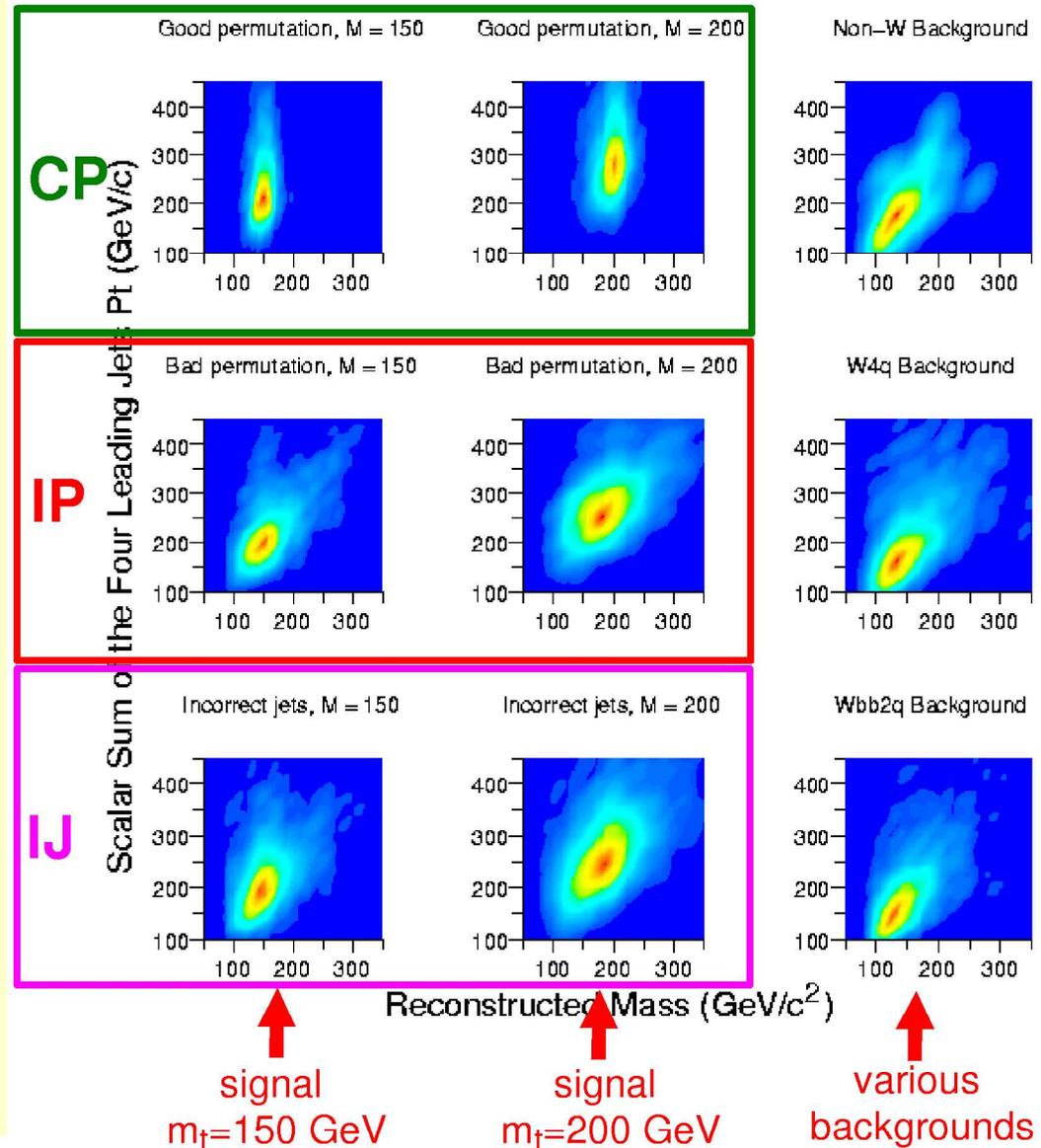


## 3) Multivariate Templates:

- In addition to the reconstructed top variables are used to improve S/B separation (avoids hard cuts)
- Sum  $p_T$  of 4 leading jets favored by statistical divergence measures

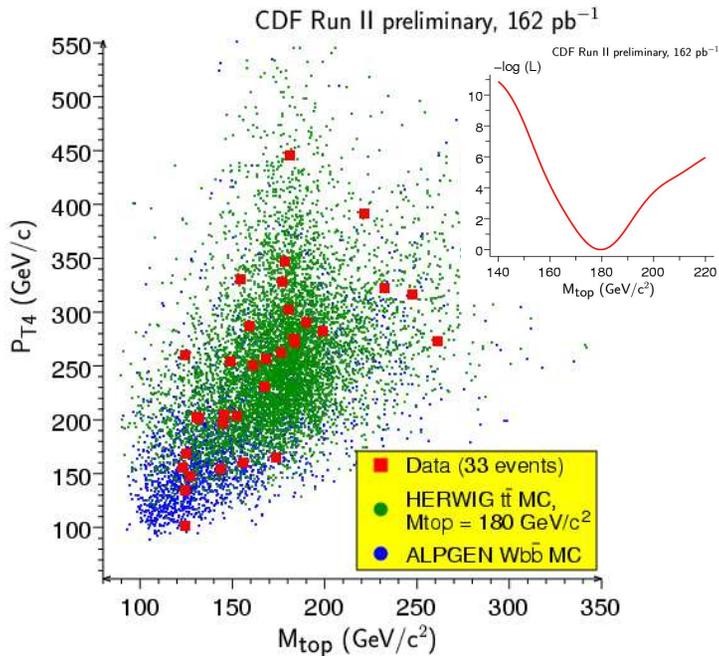


- Kernel Density Estimation is used to create probability densities (non-parametric density reconstruction technique)



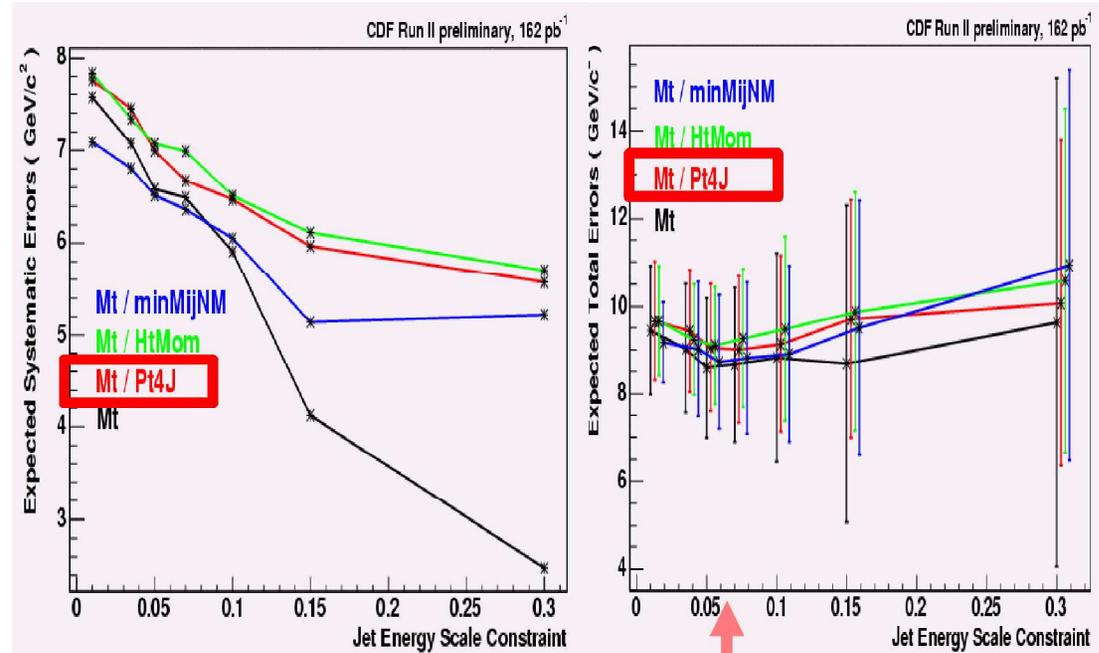


# MTM Results



Expected systematic errors:

Expected total errors:



33 candidate events

$$M_{top} = 179.6^{+6.4}_{-6.3} \text{ (stat.)} \pm 6.8 \text{ (syst.) GeV/c}^2$$

Fitted background fraction:

$$f = 0.34 \pm 0.14$$

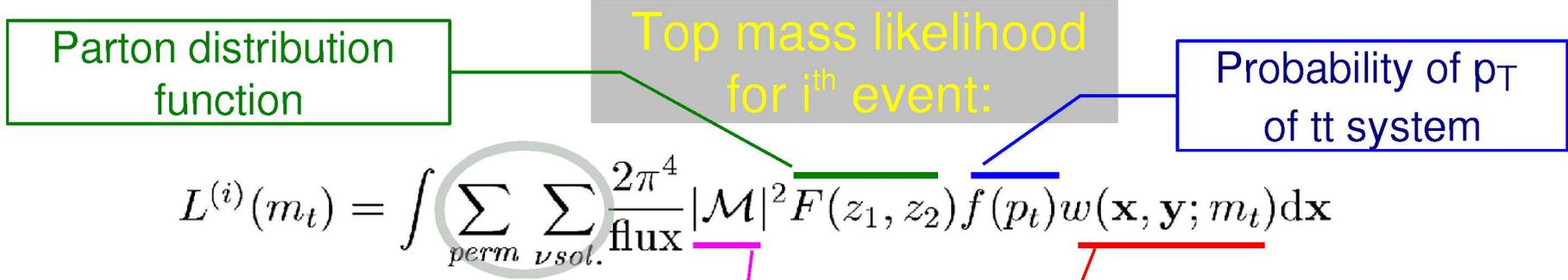
- JES constraint in kinematic fit can be used to treat with systematic errors
- JES constraint optimized w.r.t. total error
- Several variable sets with similar performance; use the one with best S/B discrimination
- JES constraint more important in the future as statistical error decreases



# Dynamical Likelihood Method (DLM)



DLM is **original CDF method** (K. Kondo, J. Phys. Soc. 57 (1988) 4126)  
It attempts to use as much amount of info on top quarks provided by SM

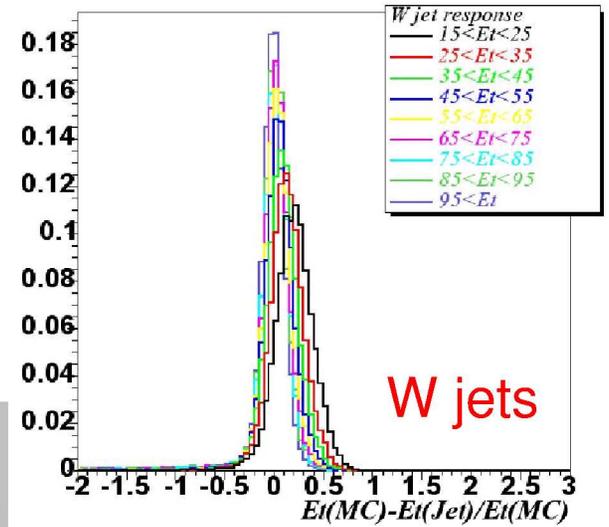


Production/Decay Matrix Element  
Signal only, no bkg M.E. ... correct for this!

Transfer function for jet energies:  
Bayesian probability that  $x$  was generated when  $y$  was reconstructed (derived from  $t\bar{t}$  MC)

- are expressed as function of  $(E(parton) - E(jet)) / E(parton)$
- are dependent on jet type (W or b jet)
- are parametrized in jet  $E_T$  and  $\eta$  bins

w jet response comparison( $E_t$ )



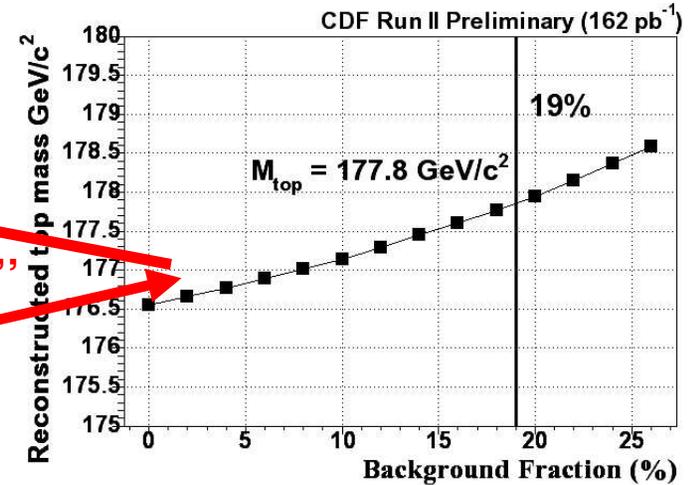
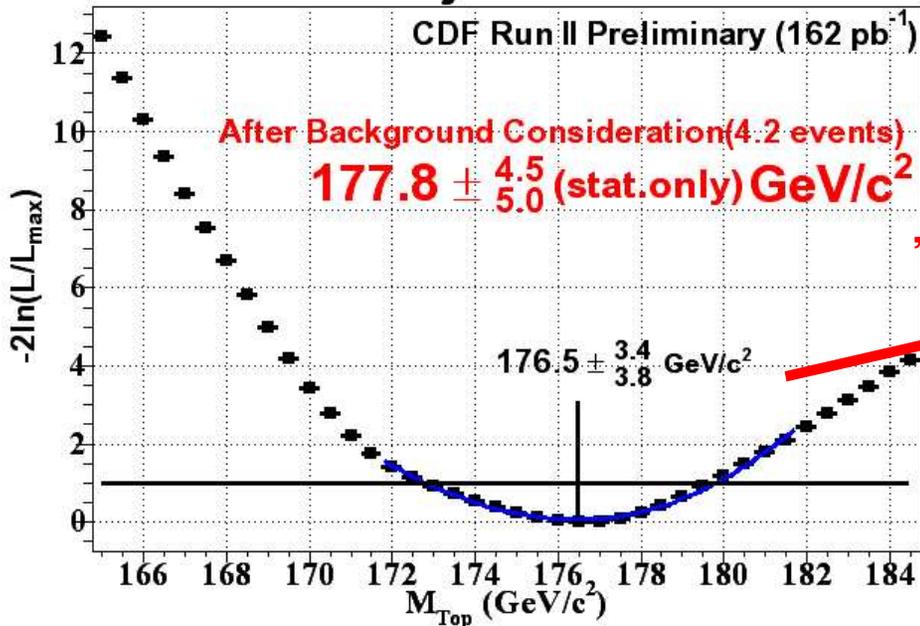
To obtain top mass, maximize  $\prod_i L^{(i)}(m_t)$



# DLM Results



## 22 events joint likelihood

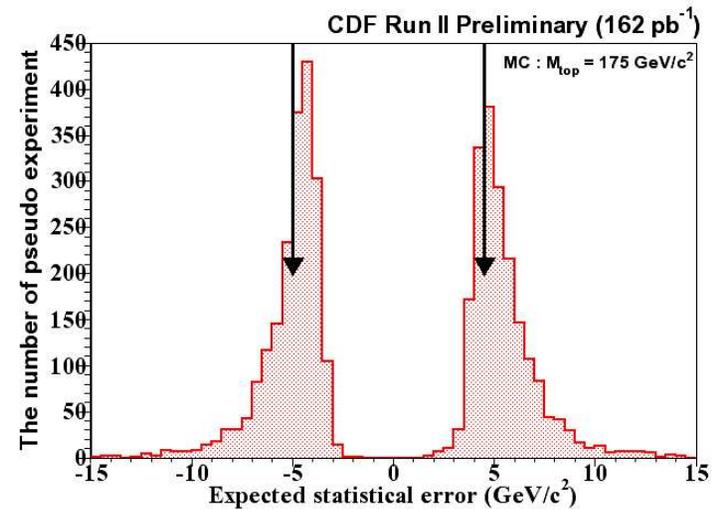


## 22 candidate events

Background fraction : 19%

After applying mapping function  
(errors scaled accordingly):

$$M_{\text{top}} = 177.8^{+4.5}_{-5.0} \text{ (stat.)} \pm 6.2 \text{ (syst.) GeV/c}^2$$

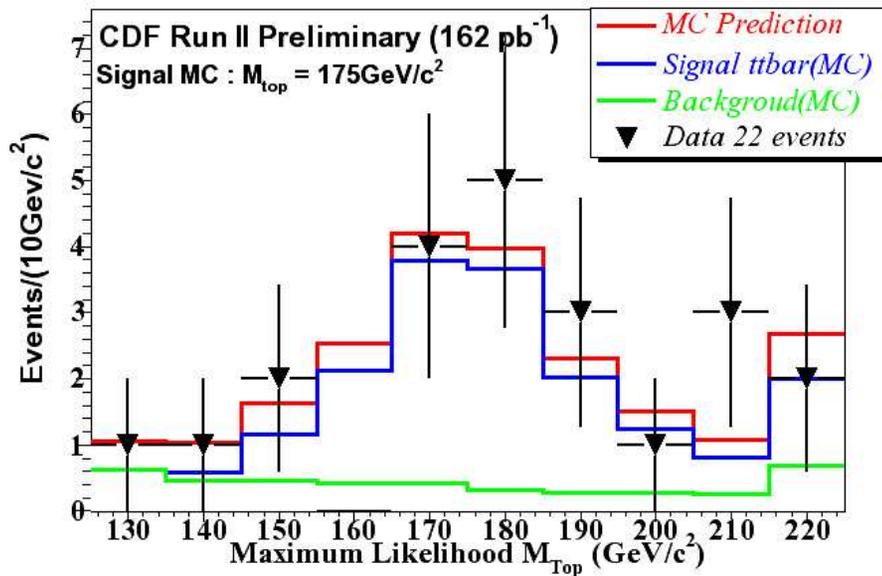


## Event-by-Event maximum likelihood mass

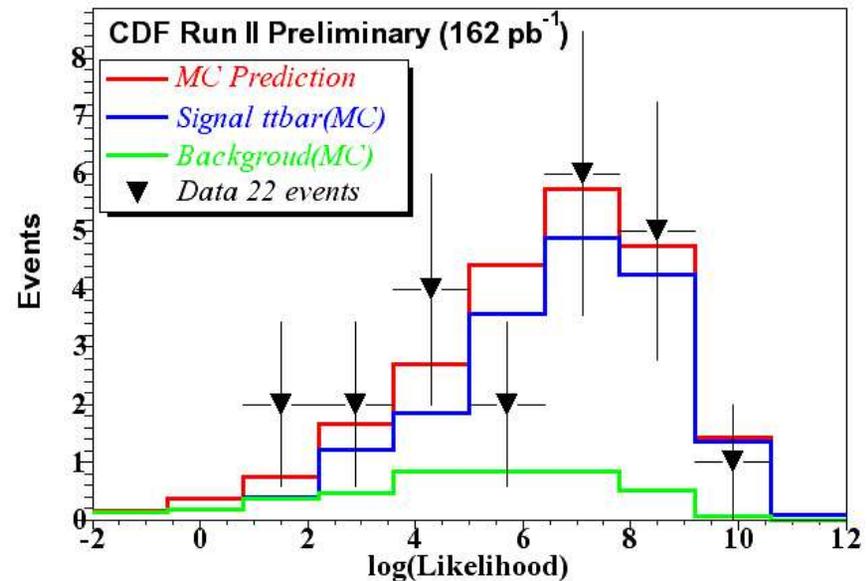
## Event Likelihood

$$L^{(i)} = \int L^{(i)}(M) dM$$

Maximum Likelihood Mass



Event Likelihood distribution



Comparison data / MC ok!



# Summary and Outlook



- Three recent CDF II measurements of the top mass in the lepton+jets channel:

**DLM method provides smallest error (attempts to use max. theory information)**

- Future improvements

★ **More data** (Tevatron performing well!):

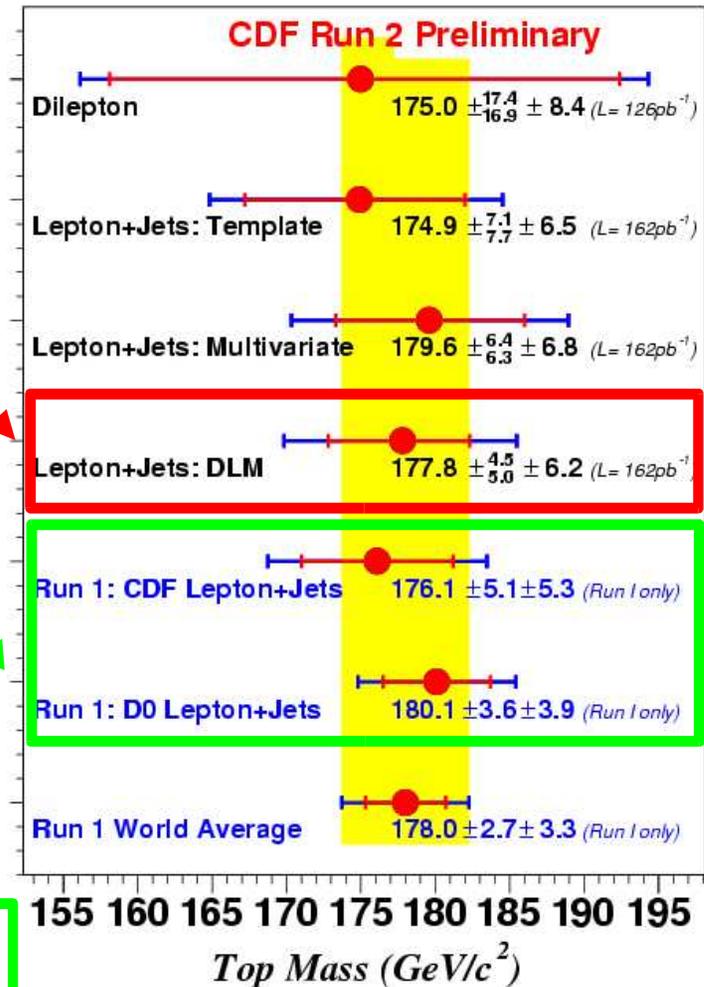
160 pb <sup>-1</sup>	present results
400 pb <sup>-1</sup>	end 2004
1000 pb <sup>-1</sup>	end 2005
4400–8500 pb <sup>-1</sup>	end Run II

★ **Reduce jet energy systematics (priority!)**

Expect significant improvement in calorimeter simulation!

★ **Big Potential to refine/extend analysis methods**

Best single measurements from Run I lepton + jets



**Precision measurement is coming ...**

**Backup Slides**



# CDF II Detector



CDF underwent substantial upgrades:

Improved geometrical acceptance:

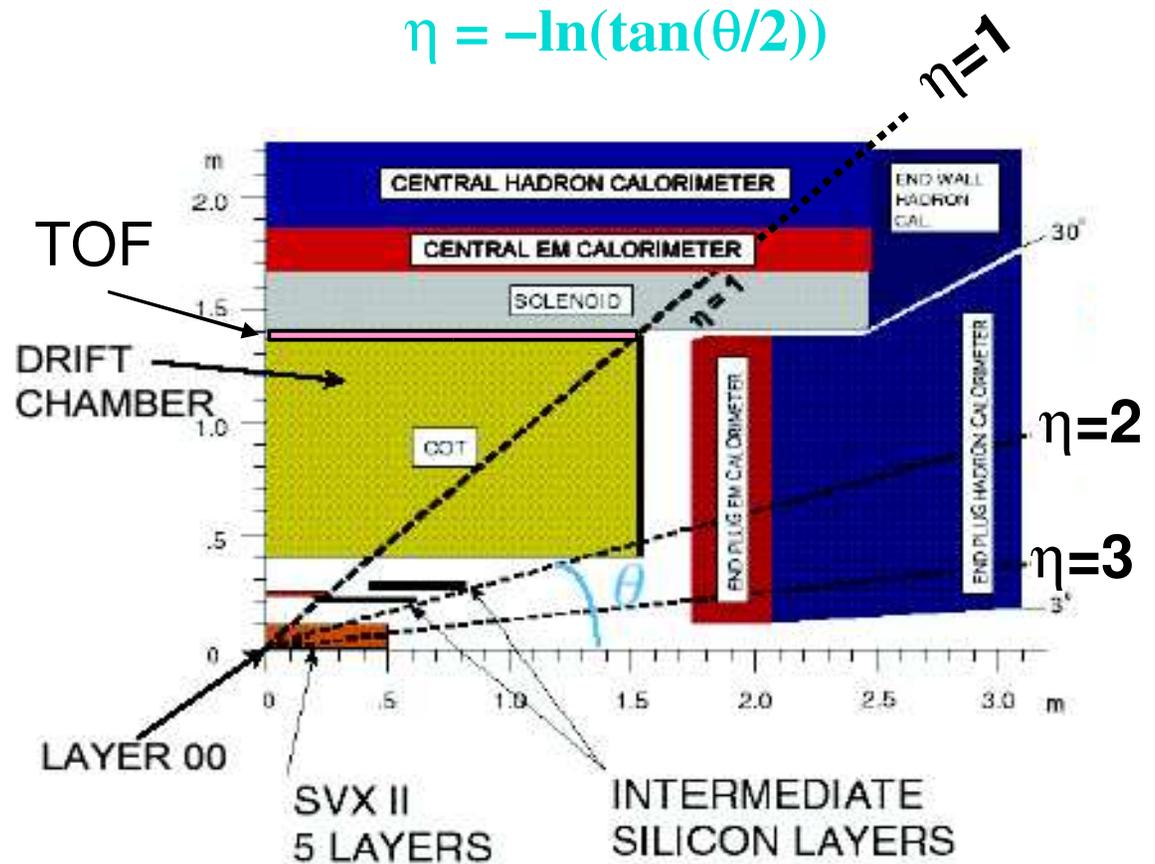
- SVX coverage  $|\eta| < 2$ , 8 layers
- Expanded Muon system
- Forward calorimeter

New central tracker

- 96 layers

Time of Flight

Trigger, DAQ ...





# MTM Correct Permutation Probability



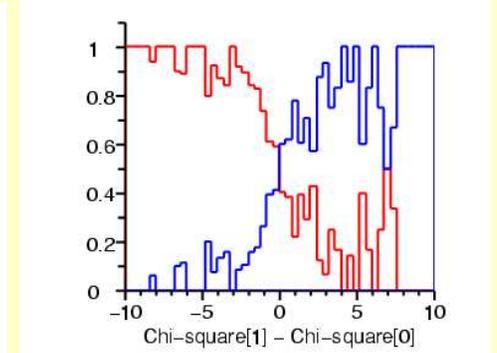
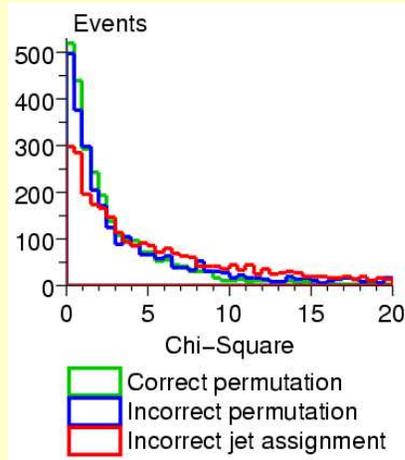
CP probability:

- Fit  $\chi^2(\text{best})$  not useful quantity
- $\chi^2(i) - \chi^2(\text{best})$  does better job

Use a "permutation diffusion" inspired model:

- 2 permutation case:

$$p_{CP} \propto 1/\exp(\chi^2(\text{2nd best}) - \chi^2(\text{best}))$$

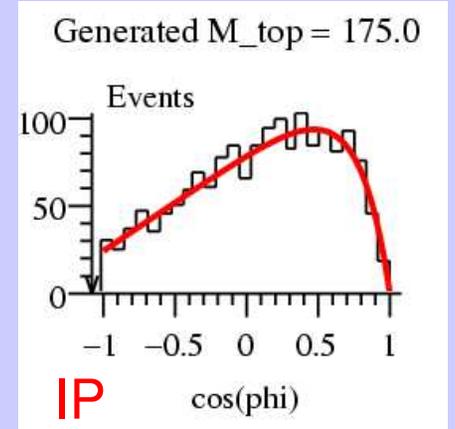
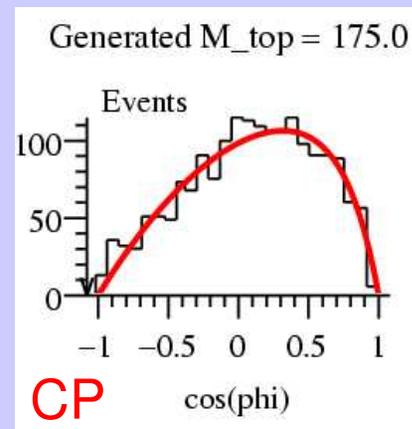


Blue: perm. 0 is correct  
Red: perm. 1 is correct

Enhance approach by adding kinematic information

- $\cos\angle(\text{lepton}, \text{leptonic } b)$
- $tt$  spin correlation term
- Define discriminator  $\kappa = p_{X|CP} / p_{X|IP+IJ}$  to be calculated using MC
- Use Bayes Statistical techniques to derive the probabilities

$$p_{CP|X} = p(p_{CP}, \kappa)$$





# MTM Correct Permutation Probability (2)



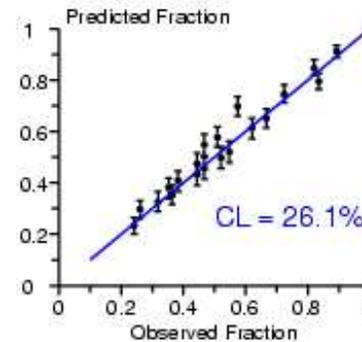
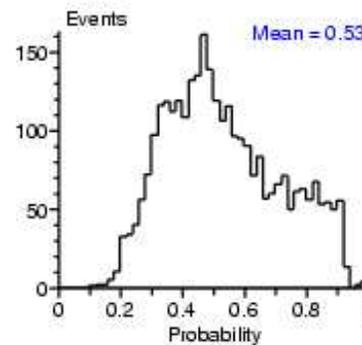
$$p_{cp} = \frac{a_b}{\sum_{i=1}^{12} a_i \exp\left(-\frac{\chi_i^2 - \chi_b^2}{w_e(\chi_i^2 + \chi_b^2)}\right)}$$

$$w_e(y) = \exp(b_e + c_e y + d_e y^2)$$

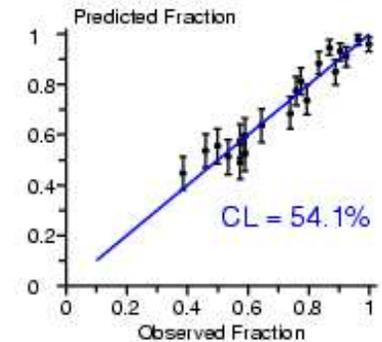
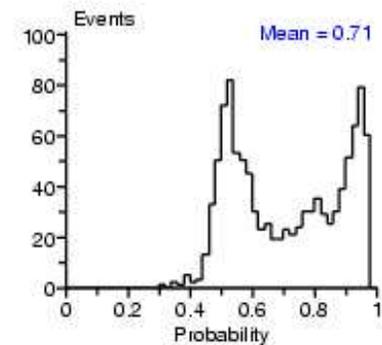
$$p_{cp|X} = \frac{\kappa p_{cp}}{\kappa p_{cp} + (1 - p_{cp})}, \quad \kappa \sim \frac{P_{X|cp}}{P_{X|\bar{c}p}}$$

Derived from MC

1 b tag



2 b tags





# MTM Likelihood



$$L(m_t) = \prod_{i=1}^N (f_b P_b(m_i, x_i) + (1 - f_b) P_s(m_i, x_i, m_t))$$

$$P_b(m, x) = \sum_{\text{bg types}} a_j B_j(m, x), \quad \sum_{\text{bg types}} a_j = 1$$

$$P_s(m, x, m_t) = (p_{cp} S_{0,m_t}(m, x) + (1 - p_{cp}) S_{1,m_t}(m, x)) p_{cj} + (1 - p_{cj}) S_{2,m_t}(m, x)$$

$N$  is the number of observed events

$m_i$  is the top mass in the  $i$ th event.  $x_i$  symbolizes all other template variables.

$P_s$  and  $P_b$  are the signal and background densities

$f_b$  is the background fraction, treated as a nuisance parameter

$B_j$  are the templates for different background types.  $a_j$  are the background composition coefficients (assumed known).

$S_{0,m_t}$ ,  $S_{1,m_t}$ , and  $S_{2,m_t}$  are the three signal templates for the given generated  $m_t$

$p_{cj}$ ,  $p_{cp}$  are the probabilities of correct jet and permutation assignments, respectively



# Systematic Uncertainties



CDF Run II Preliminary (162 pb<sup>-1</sup>)

Source	Uncertainty (GeV/c <sup>2</sup> )	
	≥ 3.5 jets	≥ 4 jets
Statistical	+7.1 / - 7.7	±6.6
<b>Systematic Jet Energy Scale</b>	<b>6.3</b>	<b>6.6</b>
Initial State Radiation	0.4	0.6
Final State Radiation	0.9	1.0
Parton Distribution Functions	0.2	0.2
Generators	0.4	0.4
Other MC Modeling	0.7	0.7
Background Shape	0.8	0.8
B-tagging	0.1	0.1
Total	6.5	6.8

CDF Run II Preliminary

Source	Uncertainty (GeV/c <sup>2</sup> )	
	≥ 3.5 jets	≥ 4 jets
Relative to Central	3.0	3.2
Central Calorimeter Response	4.6	4.7
Corrections to Hadrons (Absolute Scale)	2.2	2.3
Corrections to Partons (Out-of-Cone)	2.3	2.3
Total	6.3	6.6

I. Template Method

III. Dynamical Likelihood M.

Source	$\Delta M_{top}$ GeV/c <sup>2</sup>
<b>Jet Energy Corrections</b>	<b>5.3</b>
ISR	0.5
FSR	0.5
<b>PDFs</b>	<b>2.0</b>
Generator	0.6
Spin correlation	0.4
NLO effect	0.4
<b>Transfer Function</b>	<b>2.0</b>
Background fraction(±5%)	0.5
Background modeling	0.5
Monte Carlo modeling	0.6
Total	6.2

II. Multivariate Template Method

Systematic	$\Delta M_{top}$ (GeV/c <sup>2</sup> )
<b>Jet Energy</b>	<b>6.7</b>
Generators	0.2
ISR	0.2
FSR	0.6
PDF	0.6
Background Shape	0.4
<i>b</i> Tagging	0.3
Fitting Procedure	0.7
Total	6.8



# DLM Mapping Function

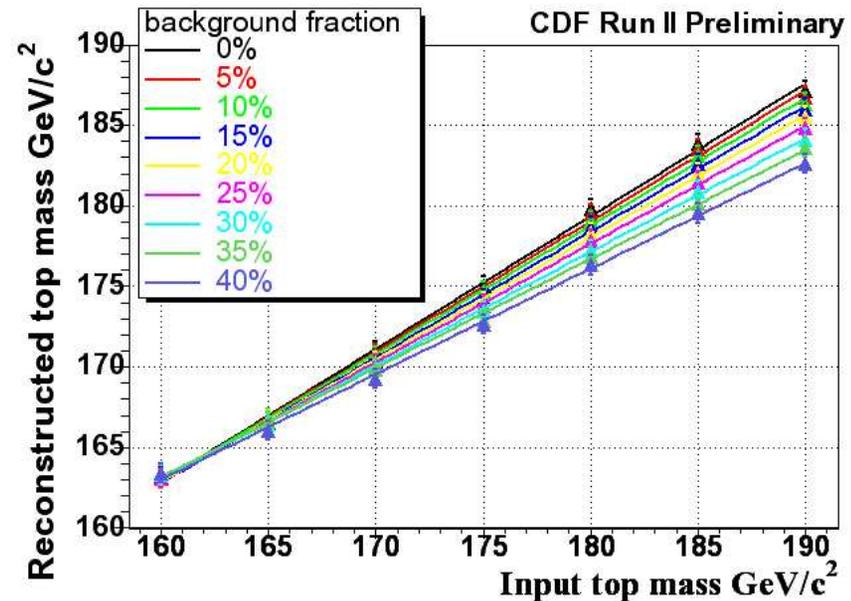


Mapping function:  
reconstructed top mass  
→ generated top mass

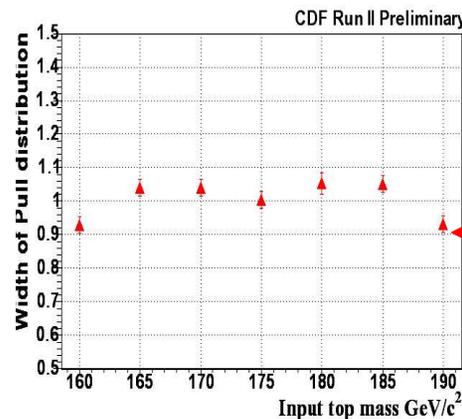
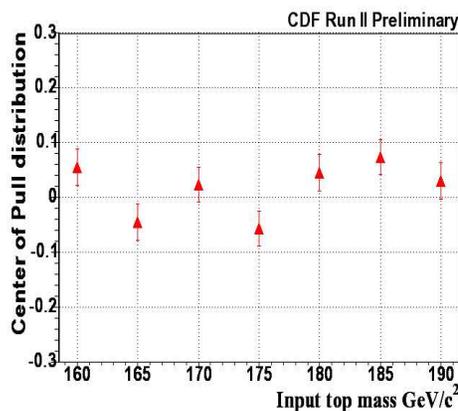
Needed to handle

- Background effects
- Top mass dependence of transfer function

Background fraction is minimized by requiring exactly 4 jets



- Slope depends on background fraction
- Expected background from cross section measurements used



Checks with pseudo experiments:

- Pulls  $\sim 0$
- Pull widths are unit Gaussian